Astro_98 Final Project Report - Griffin Beckman

Introduction:

The goal of this project was to use a small data set of galaxies in order to determine the average distance away from us for this data set. I also downloaded images of these galaxies and printed them. These are not necessarily all galaxies, some of them could be stars of other objects, but I was not able to determine which objects are galaxies or not and how to exclude them. From a small collection of galaxies, I compared their distances away from us with the given masses of the object.

Data Source:

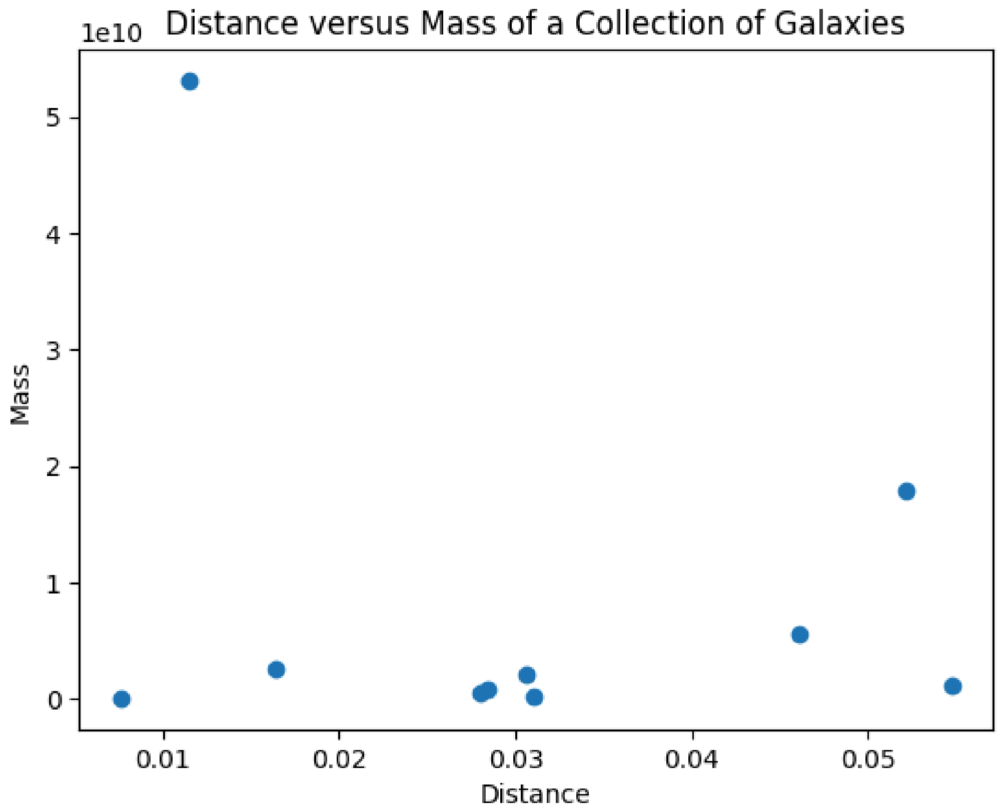
I used the fits file “nsa_v0_1_2.fits”, and used the legacy sky survey (https://www.legacysurvey.org/) in order to download the images of the galaxies that I observed. I also frequently use this website, as it has functionality that allows me to look at a random celestial object in the sky in detail, and has led to me finding some cool galaxies, etc.

Data Filtering:

I first needed to transform the fits file into a readable form, so I used the Table function form astropy to easily read the fits file. I then printed all of the different columns of the array, so that I could determine which columns would be important for my project. The first two that I determined that I needed to use were RA and DEC, which would be used for finding a given galaxy in the array on legacy survey. I also used Mass and ZDIST in order to collect the actual data that I would be using. I loaded these columns as arrays through the fits function from Astropy, and the open function from Numpy. I also printed these values to see that the values and lengths of the arrays were accurate (which they seemed to be). I then created a list of ten randomly selected integers from any index of the RA array using the numpy random function.

Data Analysis:

From here I downloaded images from a url of legacy survey, by making a function that would input a given RA and DEC value into a url, and would save them in a specific fits file on my desktop. I then opened these fits files by looping through the random integer list and opening the file with the astropy fits function and plotted them with the matplotlib imshow function. This gave me the ten images of the randomly selected galaxies. From here, I downloaded the MASS and ZDIST columns and made a list using each of the random integers as index values. From here, I averaged these values out by adding each value in the list and dividing by the length of the list. I then graphed each distance value (x-axis) against each mass value (y-axis) using the matplotlib scatter plot function. Once I had completed this I created the following graph:



My analysis of this graph is that there is relatively little, if any correlation between mass and distance, which should be expected as masses of galaxies are not generally a function of distance. However, my assumption of the data was that less massive objects would be closer to us, because the further away an object is, the smaller it appears, so they would be harder to detect, however I did not have a large enough dataset to determine if this hypothesis had any relevance. There could be a small correlation, because if we exclude the one outlying galaxy (the massive and close to us object), then galaxies tend to get slightly larger the further away they are from us.

Conclusion:
This project has not given any definite conclusion, but does suggest that for the most common range of masses, distance is weakly linearly proportional.

Sources:
I used numpy, astropy, and matplotlib documentation, and also found a research paper that also used the same data set that I used, which I used to understand which columns were important for me:
(https://iopscience.iop.org/article/10.3847/1538-4357/acdc90/pdf)